IBM

How to configure Adapter Failover (Ethernet Network Interface Backup / Link Aggregation) on the POWER processor-based blades

White Paper

Trina Bunting
Power Systems Advanced Technical Support

March 2009

Revision History
1.0 - March 9, 2009
1.2 - March 23, 2009
2.0 - December 15, 2009

Table of Contents

I. Introduction

This document provides an overview of Network Interface Backup (NIB) feature of AIX® as well as how to configure and implement NIB style EtherChannel® on IBM BladeCenter® POWER processor-based blades® running AIX or Integrated Virtualization Manager (IVM), in a BladeCenter S (BC-S) and BladeCenter H (BC-H) Chassis'. We review the configuration of the Advanced Management Module (AMM) and the Ethernet Switch Modules (ESM) including the Cisco Systems® and Blade Network Technology® (BNT, formally named Nortel Ethernet Switch Modules) It covers how to successfully test Adapter Failover from AIX, Ethernet Switch Modules (ESM) which include BNT (formally Nortel) Switches, from Blade Network Technology and Cisco Switches from Cisco Systems, as well as troubleshoot your NIB configuration. It also covers how to configure Link Aggregation (LA) in the BC-S Chassis.

II. Prerequisites

The following hardware/software was used for my NIB/LA configuration and test Environment:
- Two Ethernet Switch Modules in a BladeCenter Chassis (BNT and Cisco Systems)
- One JS12 or JS22 POWER blade
- AIX V5.3 TL7 and V6.1 TL1 were tested in this document however other levels are supported. Refer to the AIX Documentation to ensure support for the level you are using
- Integrated Virtualization Manager V2.1 FP 20.1

III. Adapter Failover (Network Interface Backup and Link Aggregation) Overview

Network Interface Backup (NIB) and Link Aggregation (LA) are features of the AIX Operating System that provide adapter failover support on the POWER processor-based blades. NIB is designed to protect against a single point of network failure by providing failure detection and failover. The goal is to make sure that a single failure (at an adapter, node, or network component) doesn't impact all the links configured for the NIB group.   The NIB/LA functions are transparent to the network equipment therefore no special hardware or network equipment is needed.

In a NIB environment, only one channel is active at any given time. When a failure is detected, all traffic is sent and received on the backup adapter (the MAC address assigned to ent0 is also used on the EtherChannel pseudo-device ent2; both Network Interface Cards (NICs) have separate MAC addresses). NIB provides rapid detection/failover (2-4 seconds) with no disruption to currently active user connections.

POWER Blades (Integrated NICs) – BC-H, BC-HT and BC-E

Blade Server

ent2
(NIB)

172.25.254.64
00:0d:60:1e:16:c0

Active   *Standby*

Primary Adapter
00:0d:60:1e:16:c0

ent0
(phy)

ent1
(phy)

Backup Adapter

Ethernet Switch Module
Bay 1 of Chassis

Ethernet Switch Module
Bay 2 of Chassis
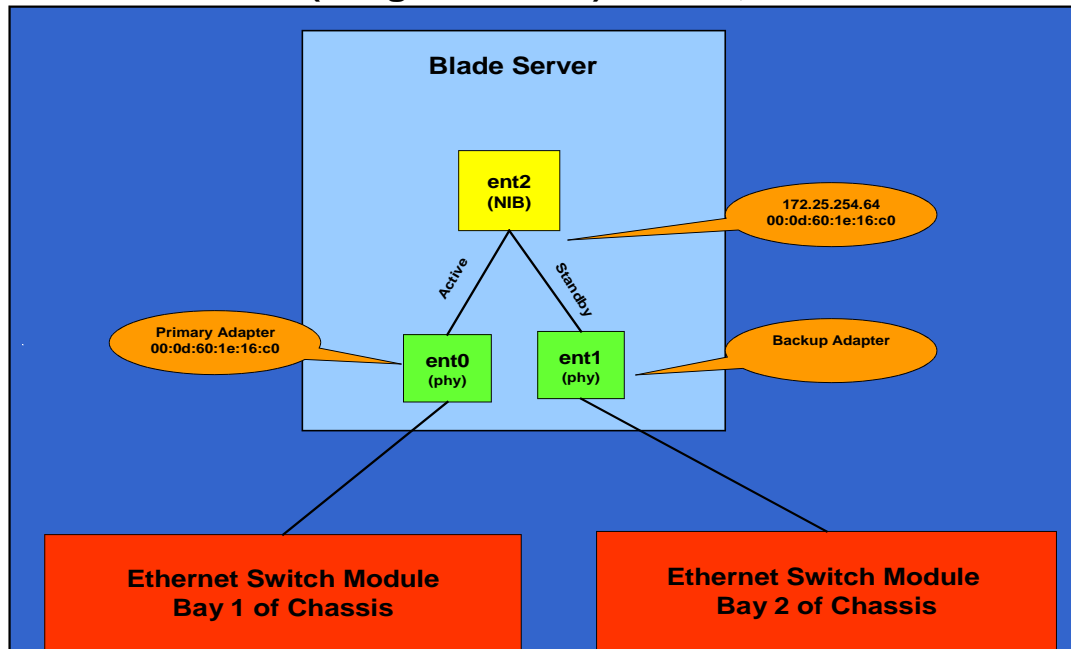
There are two key tests (link status and ping) that can be used to protect against local link failure and network path failure, respectively. When an Integrated Virtual Ethernet (IVE), also called Host Ethernet Adapter, is used as the primary in a NIB configuration the ping test should be used as the failover determination. The *ping test* checks the path through the network to the user-specified IP node. By default, the ping test sends a ping every second. If at least one response in three pings (default; can be modified by the user) is received, the network path is considered to be good. If three pings go without a response, failover is triggered.

If the ping test on ent0 fails, the NIB/LA software begins receiving traffic for the ent2 pseudo-device on the backup adapter ent1. Once the problem on ent0 has been resolved, all traffic is once again sent and received on the primary adapter. The following diagram is a sample NIB configuration where the Integrated Ethernet adapter (ent0 and ent1) is configured as EtherChannel pseudo-device ent2. The en2 interface should be configured with an IP Address, not the en0 or en1 interface. The ent0 (primary) adapter is active and ent1 is the backup adapter.

Note: If the ping on the primary adapter continues to fail, the traffic will continue to flow over the backup adapter until it itself, fails. The overall effectiveness of this solution depends upon the implementer's knowledge of network design.

The *Lossless Failover After Ping Failure* is a mechanism that ensures your backup adapter is working and available to take over traffic in case of a failover. This is achieved by sending a gratuitous arp from the primary adapter to the external core switch which routes the gratuitous arp back to the backup adapter. When the backup adapter

receives the packet then this indicates the backup adapter is healthy and the primary channel can failover. The EtherChannel pseudo-device waits for the gratuitous arps to be received via the new active channel before it finishes failing over. This is done to ensure the new active channel is working. The Lossless Failover feature is currently turned on by default. When the Lossless Failover feature is turned off the EtherChannel pseudo-device does not wait for gratuitous arp to be received via the new active channel and failover will occur rapidly. The following diagram is an example of a dual core architecture (two external core switches) in which the Lossless Failover feature should be turned off.

Dual Core Architecture - Lossless Failover = off
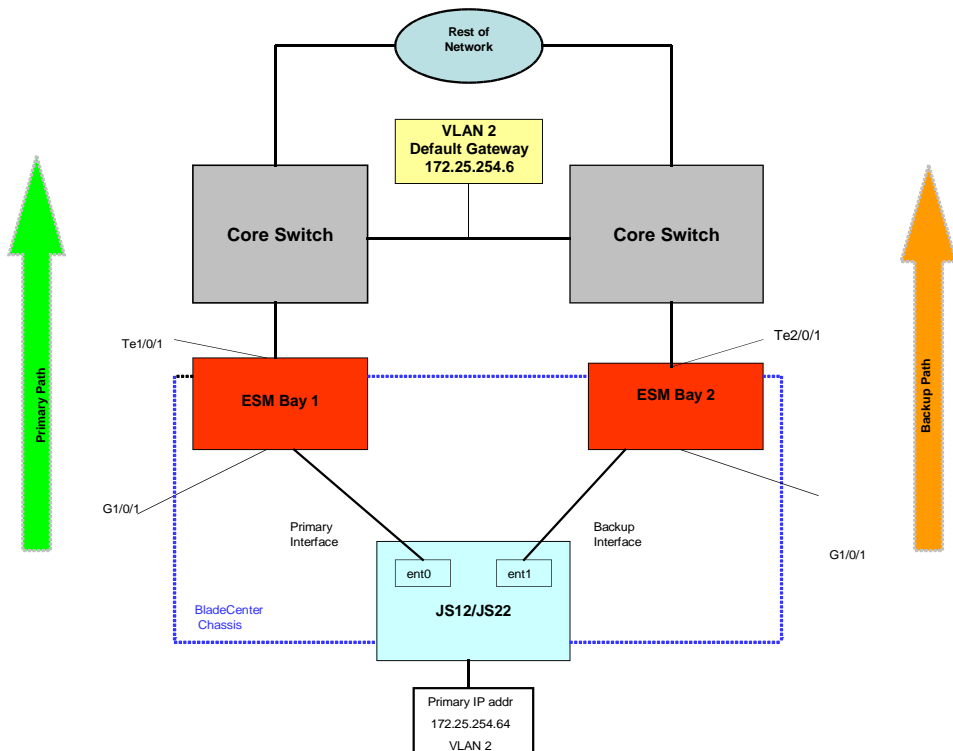
The diagram below is an example of a single core architecture (one external core switch) in which the Lossless Failover feature can be turned on (which is the default) or off.

Single Core Architecture - Lossless Failover = on/off
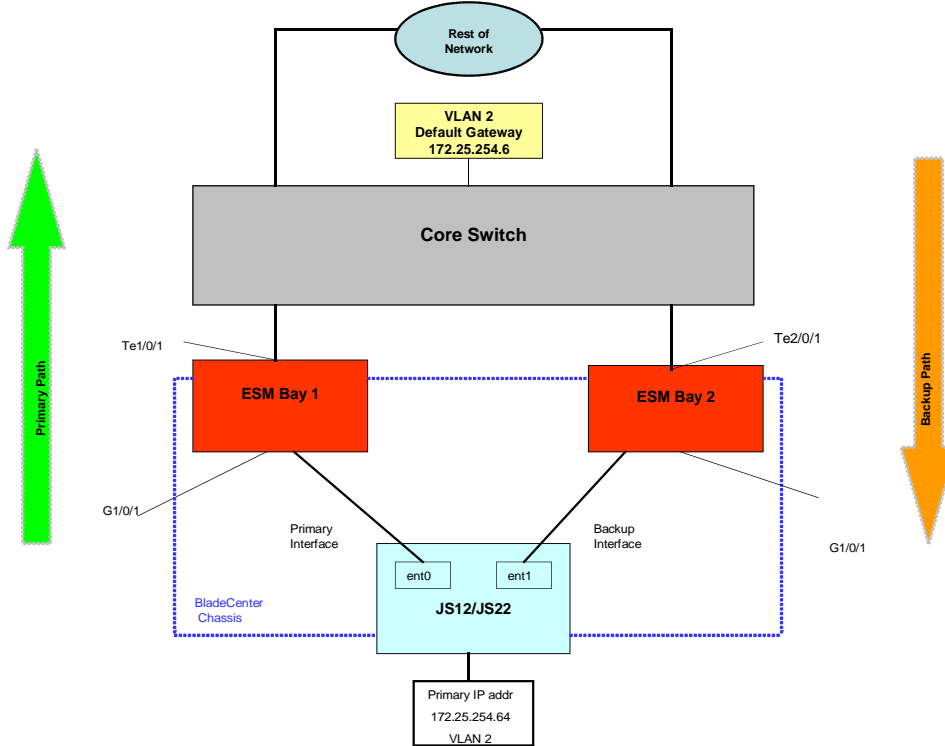
A. Configuring NIB/LA on the POWER processor-based blades

The Network Interface Backup and Link Aggregation features of AIX are supported on the POWER processor-based blades. NIB has to be configured on the POWER processor-based blades in a BC-H, BC-HT, and BC-E (JS12 only) because of how the NICs are wired in the I/O Module Bays. The integrated (NICs) on the POWER processor-based blades are hard-wired to two different Switch Module Bays (ent0 = Bay 1 and ent1 = Bay 2) in all of the BladeCenter chassis except for the BladeCenter S in which both integrated NICs are hard-wired to Bay 1 therefore Link Aggregation must be configured. As a result of this configuration, only a NIB-style EtherChannel configuration can be implemented on those chassis (BC-H, BC-HT, BC-E).   In a NIB configuration ent0 should be configured as the primary adapter (Active) and ent1 should be configured as the backup adapter (Standby).

In addition to the integrated NIC on the POWER processor-based blades, an Ethernet expansion card (CFFv and/or CFFh) can be used for additional Ethernet ports. The port on the CFFv card is hard-wired to Bays 3 and 4 of the chassis. The ports on the CFFh card are hard-wired to Bays 7-10 of the BC-H and BC-HT (except for the BC-S where the ports are hard-wired to Bay 2) therefore an MSIM or MSIM-HT is required.

B. POWER processor-based blades NIB/LA Configuration Rules

1. When configuring EtherChannel with IVE, all primary adapters should be of the same type (i.e. IVE). However, the backup adapter can be a virtual adapter or a physical NIC.

2. The SoL function will only work on the primary adapter (ent0) therefore; this function will not be available when the backup adapter (ent1) is activated. During this time, the administrator would use telnet to access the blade.

3. The link state cannot be used as the failover determination because the IVE adapter on the JS12/JS22 blades will **always show a link status of Up**:

   $ **entstat –d ent# | grep Link** or **entstat –all ent# | grep Link** (VIOS)
   Logical Port Link State: Up
   Physical Port Link State: Up

   Note: The link state change on blades is not propagated to the IVE driver in AIX because there is a switch which sits between the IVE coming off the P5IOC (blade I/O controller) and the serializer deserializer (SerDes).

4. In a NIB configuration EtherChannel should NOT be configured on the downstream ports but in a LA configuration EtherChannel MUST be configured on the downstream ports.

5. The Lossless Failover feature in NIB should be turned off if the customer's BladeCenter is configured in dual core architecture (two external core switches).

6. EtherChannel Teaming (LA) is not currently supported in the BC-H, BC-HT and BC-E chassis when the BladeCenter switches are configured in stacked mode.

IV. Configuring the AMM and Ethernet Switch Modules in a Failover Environment

It is important to ensure the AMM and the ESM are properly configured for failover environment in the chassis as most of the chassis support NIB and only one supports LA. NIB is supported when the POWER processor-based blades are running in the BC-H,

BC-HT and BC-E Chassis and LA is supported in the BC-S Chassis. The following features must be configured from the AMM:

1. Ensure the External Ports for each I/O Module is "**Enabled**" by going to "**Admin/Power/Restart**" from I/O Module Task and select I/O Module Advanced Setup.

2. It is important to consider how the ESM is managed before implementing NIB into your environment. If the user wants to manage ESM from the AMM, the "**External management over all ports**" to should be set to "**Disabled**" in the MM Advanced Management section for each ESM).   If the user wants to manage the Switch via its uplink ports then this feature should be enabled.

3. The External Network Interface (eth0) from the AMM to the Distribution Layer Switch must be configured. The Advanced Ethernet Setup for eth0 should have failover "**Disabled**".

Although the default ESM configuration should work in a NIB environment, most customers inevitably customize their configuration. If your switch modules have been

customized and you are having trouble with NIB, refer to the *Trouble Shooting* Section of this document.

A. Configuring Link Aggregation in a BladeCenter S Chassis

In a BladeCenter S (BC-S) Chassis, Link Aggregation (LA) is used for adapter failover instead of NIB because the integrated NICs on the blades are hard-wired to the same switch module (Bay 1). An Ethernet expansion card is required on the POWER processor-based blades (CFFh, feature code 8291 or part number 44W4479) as the NICs on that card are hard-wired to Bay 2 of the chassis. It is important to note that EtherChannel must be configured on the downstream ports via the ESM when configuring Link Aggregation. The following diagram is an example of Link Aggregation configured on a blade running AIX with an Ethernet Expansion Card (CFFh) installed:

POWER Blade/CFFh Exp Card – BC-S
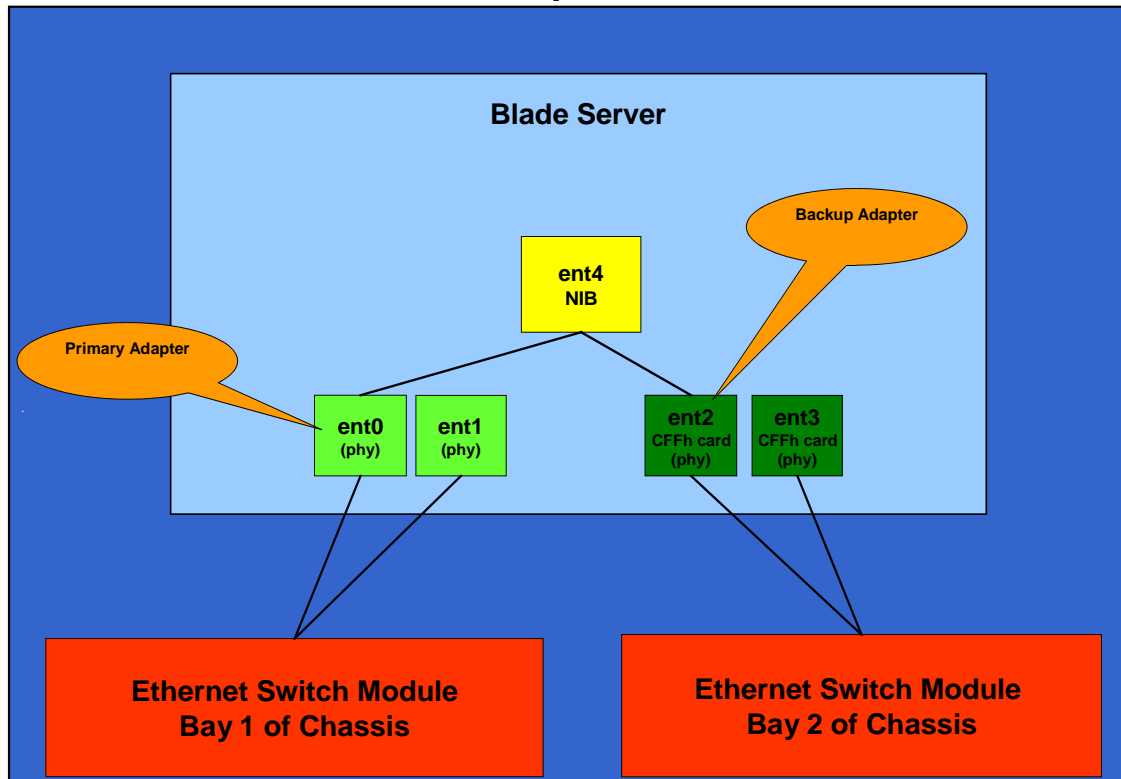
Note: An RPQ (8A1747) is required to enable the 1 GB Ethernet/ 4Gb Fibre Channel Combo Expansion Card (CFFh), feature code #8252 installed on the blade to connect to the Ethernet Switch Module in bay 2 of the BC-S chassis. The Fibre Channel connections are not enabled in the BladeCenter S chassis with this RPQ.

B. Cisco Systems Ethernet Switch Modules in a NIB/LA Environment

Cisco Systems produces a lot of different Ethernet Switch Modules including:

- Cisco Catalyst Switch Module 3110x(10GB uplink)
- Cisco Catalyst Switch Module 3110g (1 GB Uplink)
- Cisco Catalyst Switch Module 3012
- Cisco Systems Intelligent Gigabit Ethernet Switch Module
- Cisco Systems Fiber Intelligent Gigabit Ethernet Switch Module

The default configuration on all of these switches except the Cisco Systems Intelligent Gigabit Ethernet Switch Module (CIGESM) will allow deployment of NIB without any need for configuration. When using the CIGESM it is important to hard code the speed of the ports going to the JSXX blades on both CIGESM. Telnet to the Cisco Switch Module and type the following commands on both switches:
1. Type "**config t**" and press "**Enter**"
2. Type "**int range g0/1 -14**" and press "**Enter**"
3. Type "**speed noneg**" and press "**Enter**"
4. Type "**end**" and press "**Enter**"
5. Type "**write**" and press "**Enter**"

For more advanced configuration options, refer to the documentation on the 3110 and 3112 Cisco Catalyst Switches at:
http://www-947.ibm.com/systems/support/supportsite.wss/docdisplay?lndocid=MIGR-5075938&brandind=5000020 and the Cisco Systems Intelligent Gigabit Ethernet Switch Module for IBM eServer BladeCenter Redpaper at (http://www.redbooks.ibm.com/redpapers/pdfs/redp3869.pdf).

Note: The Cisco 3110X, 3110G and 3012 switches are supported in the BladeCenter H, BladeCenter HT and BladeCenter E Chassis and the MSIM. The Cisco 3012 Switch is supported in the BladeCenter S Chassis. The JS22 is not supported in the BC-E Chassis but the JS12 is supported in the BC-E Chassis.

C. Blade Network Technologies Ethernet Switch Modules in a NIB/LA Environment

BNT produces a number of switches for the BladeCenter environment, including the following:

- BNT 1/10Gb Uplink Ethernet Switch Module
- BNT Layer 2/3 Copper and Fiber Gigabit Ethernet Switch Modules
- BNT Layer 2-7 Gigabit Ethernet Switch Module
- BNT 6-port 10Gb Ethernet Switch Module (only supported in high-speed switch bays)

Other than the BNT 6-port 10Gb module, all of these switches are supported in all BladeCenter environments. The default configuration of these switches as shipped will allow deployment of NIB without any need for configuration. For those desiring more advanced configuration options, various Application Guides and Command References can be found at the following link for all of the BNT switches:
http://www.bladenetwork.net/IBM-BladeCenter-support.html.

V. Sample POWER blade NIB/LA Environment

In this sample environment, the CIGESM are managed using the AMM's uplink port. The AMM and Data Traffic are in a common VLAN. The CIGESMs are on an internally different VLAN, but shares AMM uplink VLAN for management. SoL is provided on a separate VLAN. Using a single VLAN to carry both user and management traffic is **NOT** recommended in a Production Environment. The following is a diagram of a sample test environment:

Sample POWER Blade Environment
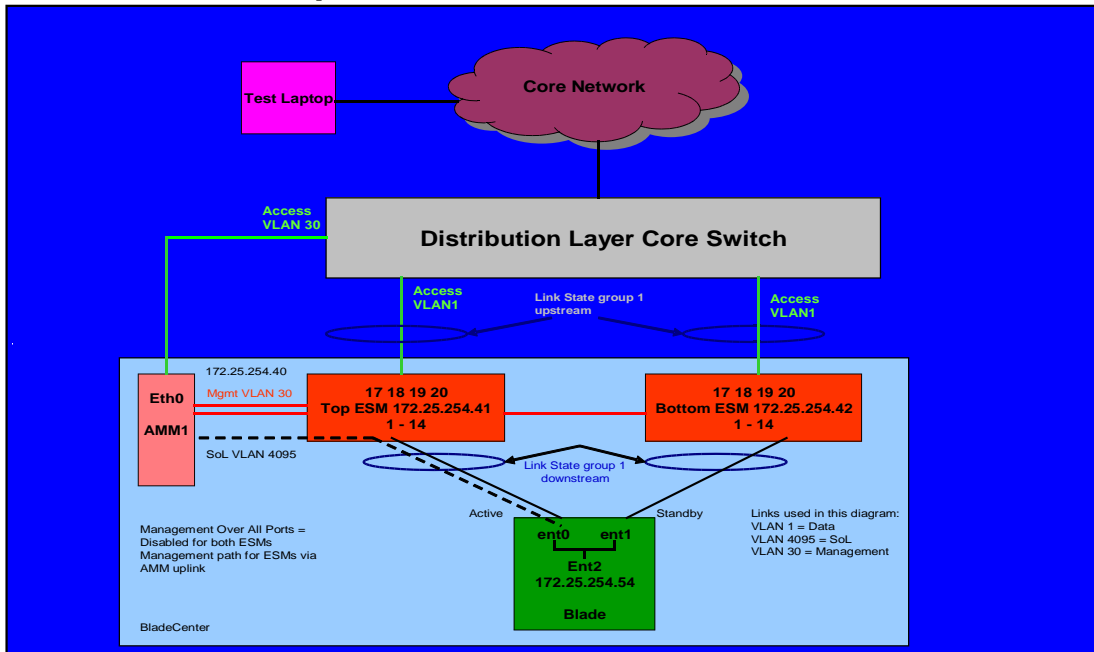
Note: The uplink ports on the ESM go upstream to the Distribution Layer Switch. The POWER blade ports go downstream from the ESM to the Blade.

In the sample Power blade environment, the VLANs are carried as follows:
- On ports going to the JS12/JS22 Blades (g0/1 – g0/14):
  - Data VLAN 1
  - SoL VLAN 4095

- On ports going to the AMM (g0/15 – g0/16):
  - AMM/Switch VLAN 30
  - SoL VLAN 4095

- On ports going to the JS12/JS22 Production Network (g0/17 – g0/20):
  - Data VLAN 1

A. Configuring NIB/LA in AIX

The following steps will configure NIB/LA on the POWER blade:
1. This step should only be performed if the adapter interfaces are configured. To remove any pre-existing IP interface configurations type the following commands:
   Type "**ifconfig en0 detach**" and press "**Enter**"
   Type "**rmdev –dl en0**" and press "**Enter**"

   Type "**ifconfig en1 detach**" and press "**Enter**"

Type "**rmdev –dl en1**" and press "**Enter**"

Type "**ifconfig et0 detach**" and press "**Enter**"
Type "**rmdev –dl et0**" and press "**Enter**"

Type "**ifconfig et1 detach**" and press "**Enter**"
Type "**rmdev –dl et1**" and press "**Enter**"

Note: The above commands can also be executed in smit from the fast path "**smitty tcpip**".

2. Verify the network interfaces have been removed:
   Type "**netstat –in**" and press "**Enter**"
   (You should only see the loopback interface)

3. Create the ent2 pseudo-device:
   a. Type the fast path "**smitty etherchannel**" and press "**Enter**"
   b. Select "**Add an EtherChannel / Link Aggregation**" and press "**Enter**"
   c. Select "**ent0**" as the primary adapter and "**ent1**" as the backup adapter
   d. The **Perform Lossless Failover After Ping Failure** defaults to **yes** but should be changed to **no** depending on the switch configuration.
   e. Enter the default gateway as the "**Internet Address to Ping**"

```
                    Add An EtherChannel / Link Aggregation

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                              [Entry Fields]
  EtherChannel / Link Aggregation Adapters    ent0                     +
  Enable Alternate Address                    no                       +
  Alternate Address                           []                       +
  Enable Gigabit Ethernet Jumbo Frames        no                       +
  Mode                                        standard                 +
  Hash Mode                                   default                  +
  Backup Adapter                                                       +
       Automatically Recover to Main Channel  yes                      +
       Perform Lossless Failover After Ping Failure  yes               +
  Internet Address to Ping                    []
  Number of Retries                           []                      +#
  Retry Timeout (sec)                         []                      +#




Esc+1=Help           Esc+2=Refresh         Esc+3=Cancel          Esc+4=List
Esc+5=Reset          Esc+6=Command         Esc+7=Edit            Esc+8=Image
Esc+9=Shell          Esc+0=Exit            Enter=Do
```

Note: The *Automatically Recover to Main Channel* feature is ignored if failover has occurred due to ping test. If the BladeCenter is configured in a dual core architecture (two external core switches) the *Lossless Failover After Ping Failure* feature should be turned off.

4. Configure the IP Address for the pseudo-device ent2:
   a. Type the fast path "**smitty chinet**" and press "**Enter**"
   b. Select "**en2**" interface and press "**Enter**"

```
                Change / Show a Standard Ethernet Interface

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                                    [Entry Fields]
    Network Interface Name                          en2
    INTERNET ADDRESS (dotted decimal)               [172.25.254.64]
    Network MASK (hexadecimal or dotted decimal)    [255.255.255.0]
    Current STATE                                   up                    +
    Use Address Resolution Protocol (ARP)?          yes                   +
    BROADCAST ADDRESS (dotted decimal)[17C[]
    Interface Specific Network Options
      ('NULL' will unset the option)
      rfc1323                                       []
      tcp_mssdflt                                   []
      tcp_nodelay                                   []
      tcp_recvspace                                 []
      tcp_sendspace                                 []
    Apply change to DATABASE only                   no                    +

Esc+1=Help          Esc+2=Refresh       Esc+3=Cancel        Esc+4=List
Esc+5=Reset         F6=Command          F7=Edit             F8=Image
F9=Shell            F10=Exit            Enter=Do
```

5. Enter the "**IP/Address, Network Mask**", change the "**Current STATE**" to "**up**" and press "**Enter**"

Adapter failover can also be configured in Linux. For more information on configuring this EtherChannel Bonding from Red Hat go to http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5/html/Deployment_Guide/s2-networkscripts-interfaces-chan.html and to configure from SLES go to http://www.novell.com/support/php/search.do?cmd=displayKC&docType=ex&bbid=TSEBB_1222707479531&url=&stateId=0 0 34017274&dialogID=34013800&docTypeID=DT_TID_1_1&externalId=3929220&sliceId=2&rfId= .

VI. VIOS High Availability

For increased availability when you have only one VIOS Server it is recommended that you setup Network Interface Backup (NIB) to protect against a single point of network failure by providing failure detection and failover. If an adapter becomes unavailable, the packets are automatically sent on the next available adapter without disruption to existing connections.

A.   VIOS in a NIB/LA Environment

The following diagram is the recommended configuration when a POWER processor-based blades has NIB configured on a single VIOS server to maximize redundancy. The Host Ethernet Adapters (ent0 and ent1) on the VIOS Server are configured as EtherChannel pseudo-devices (ent6 and ent7) depending on the chassis the Blade is. The pseudo-device should be configured to run TCP/IP therefore the physical devices will not be used. The primary adapter is active and the backup adapter is not active.

POWER Blades/CFFh Exp Card – BC-H, BC-HT, BC-E
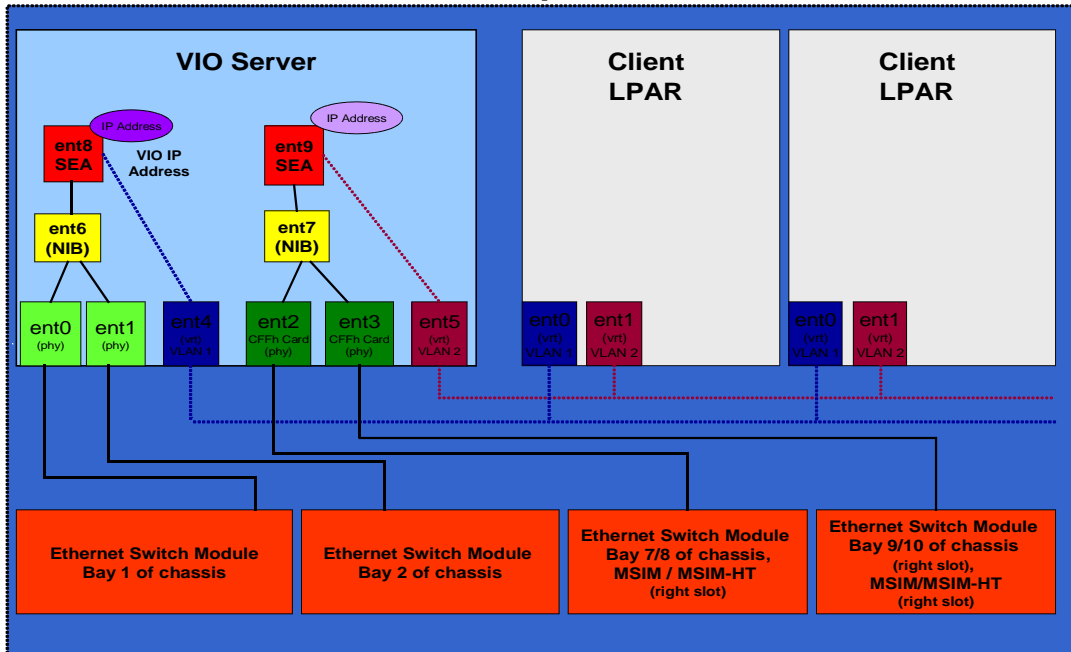
In the above diagram, the POWER blade has a CFFh Expansion Card installed which means the NICs on the expansion card are hard-wired to I/O Modules in an MSIM or MSIM-HT.

Note: The 2/4 GbE CFFh card has been withdrawn but will be introduced when we get AIX drivers (planned 10/ 2010). This card only operates as a 2 port adapter.

B. Configuring NIB/LA on VIOS

Type the following to configure NIB/Link Aggregation on the VIOS Server from the CLI:

1. Open a SoL Session to your blade: telnet to the AMM and type "**console –T blade[#]**" (where # is the Bay number of the blade) and press "**Enter**".
   a. **Note:** If you try to configure Link Aggregation from a virtual terminal and not SoL or KVM, you will lose your session.

2. From the VIOS shell type "**lsdev –type adapter**" and press "**Enter**".
3. Before you configure Link Aggregation, view the existing IP Addresses by typing "**netstat –state –num**" and press "**Enter**".
4. Remove any pre-existing IP interface configurations from the adapters by typing "**rmtcpip –all**" and press "**Enter**". Type "**y**" and press "**Enter**" to complete. To remove the shared adapters type:
   rmdev –dev et#
   rmdev –dev ent#
   rmdev –dev en#

5. To view all adapters type "**lsdev –type adapter**" and press "**Enter**".
6. To enable Ethernet Bridging (promiscuous mode) on IVE physical port 0 and 1 type the following:
   "**chhwres -r hea -o s -l 23000000 -g 1 -a promisc_lpar_id=1 --physport 0**" and press "**Enter**".
   "**chhwres -r hea -o s -l 23000000 -g 1 -a promisc_lpar_id=1 --physport 1**" and press "**Enter**".

   Note: An Ethernet Bridge allows virtual Ethernet devices to access a physical Ethernet device thereby allowing access to the external network via the physical Ethernet device.

7. To configure ent0 as the primary adapter and ent1 as the backup adapter on the VIO Server type "**mkvdev –lnagg ent0 –attr backup_adapter=ent1**" and press "**Enter**".

   Note: When configuring Network Interface Backup/Link Aggregation on a blade, you cannot mix IVE and non-IVE-adapters.

8. Type "**lsdev –dev ent6 –attr**" and press "**Enter**" to confirm the configuration of the Link Aggregation device.

```
$ chhwres -r hea -o s -l 23000000 -g 1 -a promisc_lpar_id=1 --physport 1
$ mkvdev -lnagg ent0 -attr backup_adapter=ent1
ent6 Available
en6
et6
$ lsdev -dev ent6 -attr
attribute        value        description                              user_
settable

adapter_names    ent0         EtherChannel Adapters                    True
alt_addr         0x000000000000 Alternate EtherChannel Address         True
auto_recovery    yes          Enable automatic recovery after failover True
backup_adapter   ent1         Adapter used when whole channel fails    True
hash_mode        default      Determines how outgoing adapter is chosen True
mode             standard     EtherChannel mode of operation           True
netaddr          0            Address to ping                          True
noloss_failover  yes          Enable lossless failover after ping failure True
num_retries      3            Times to retry ping before failing       True
retry_time       1            Wait time (in seconds) between pings      True
use_alt_addr     no           Enable Alternate EtherChannel Address    True
use_jumbo_frame  no           Enable Gigabit Ethernet Jumbo Frames     True
$
```

9. To create the SEA using the Link Aggregation device as the physical adapter type "**mkvdev –sea ent6 –vadapter ent3 –default ent3 –defaultid 1**" and press "**Enter".**
10. To confirm the creation of the SEA device type "**lsdev –type adapter**" and press "**Enter**".
11. To configure an IP Address on the newly created SEA type "**mktcpip –hostname (hostname) –inetaddr (IVM IP address) –interface en7 –netmask 255.255.255.0**

–gateway 172.25.254.6 –nsrvaddr 172.12.0.1 –nsrvdomain training.sc.ibm.com"
and press "**Enter**".

12. To check the IP address on the SEA type "**netstat –num –state**" and press "**Enter**".

VII. Testing Failover

The following sections show you how to test Adapter Failover and Fail back in NIB:

A. Testing Adapter Failover from AIX

To Test Adapter Failover from AIX do the following:
1. Type the fast path "**smitty etherchannel**" and press "**Enter**".
2. Select "**Force a Failover In An EtherChannel / Link Aggregation**" and press "**Enter**".

```
                    EtherChannel / IEEE 802.3ad Link Aggregation

Move cursor to desired item and press Enter.

  List All EtherChannels / Link Aggregations
  Add An EtherChannel / Link Aggregation
  Change / Show Characteristics of an EtherChannel / Link Aggregation
  Remove An EtherChannel / Link Aggregation
  Force A Failover In An EtherChannel / Link Aggregation
```
```
Esc+1=Help           Esc+2=Refresh       Esc+3=Cancel        F8=Image
F9=Shell             F10=Exit            Enter=Do
```

3. Select "**ent2**" adapter and press "**Enter**".
4. Press "**Enter**" on the pop-up menu confirming your action.
5. To determine which adapter is active type: "**netstat –v | grep Active**" and press "**Enter**".

To verify adapter failover check the error report for an error message (ECH_PING_FAIL_PRMRY) indicating the primary EtherChannel failed and the backup adapter has taken over.

```
LABEL:           ECH_PING_FAIL_PRMRY
IDENTIFIER:      9F7B0FA6

Date/Time:       Fri Jan 30 19:05:08 EST 2009
Sequence Number: 47
Machine Id:      0000354AD400
Node Id:         nfsclient
Class:           H
Type:            INFO
Resource Name:   ent6
Resource Class:  adapter
Resource Type:   ibm_ech
Location:

Description
PING TO REMOTE HOST FAILED

Probable Causes
CABLE
SWITCH
ADAPTER

Failure Causes
CABLES AND CONNECTIONS

        Recommended Actions
        CHECK CABLE AND ITS CONNECTIONS
        IF ERROR PERSISTS, REPLACE ADAPTER CARD.

Detail Data
FAILING ADAPTER
PRIMARY
SWITCHING TO ADAPTER
ent1
Unable to reach remote host through primary adapter: switching over to backup a
apter
```

If you repeat the above steps and check the error report again, an error message
(ECH_PING_FAIL_BCKUP) indicates the primary adapter has been recovered.

```
LABEL:           ECH_PING_FAIL_BCKP
IDENTIFIER:      5FC2DD4B

Date/Time:       Fri Jan 30 19:01:44 EST 2009
Sequence Number: 46
Machine Id:      0000354AD400
Node Id:         nfsclient
Class:           H
Type:            INFO
Resource Name:   ent6
Resource Class:  adapter
Resource Type:   ibm_ech
Location:

Description
PING TO REMOTE HOST FAILED

Probable Causes
CABLE
SWITCH
ADAPTER

Failure Causes
CABLES AND CONNECTIONS

        Recommended Actions
        CHECK CABLE AND ITS CONNECTIONS
        IF ERROR PERSISTS, REPLACE ADAPTER CARD.

Detail Data
FAILING ADAPTER
ent1
SWITCHING TO ADAPTER
PRIMARY
Unable to reach remote host through backup adapter: switching over to primary ad
apter
```

To determine whether the Primary or the backup adapter is active from the AIX
command type "**netstat –v | grep Active**" and press "**Enter**".

Note: Regardless of which adapter is active, ent0 will always be the primary and ent1 will always be the backup adapter.

B.   Testing Adapter Failover from the Cisco Switch Modules

There are a number of ways to test the failover feature, including physically pulling the upstream cables or disabling ports at various locations in the path between the BladeCenter chassis and the upstream switches. The following is an example of how to disable the blade facing ports on the Cisco Switch Modules.

Note: In this example both ESMs are being managed via their own uplinks. If you're testing failover by powering down the uplink ports on the ESM you will be shutting down the path by which you are managing the switch. In this case you would have to test failover by plugging directly into the console port of the ESM to test failover.

Shutdown the Blade Facing Ports on the Cisco Switch Modules

From enable mode on the Cisco switch type "**show int status**" for a list of the ports on the ESM.
1. Setup looping ping from a network attached device (i.e. your laptop) to the sever with the NIB setup (ping –t ipaddress).
2. The following commands should be typed on both ESMs, starting with the ESM 1:
    a. Type "**config t**" and press "**Enter**"
    b. Type "**int G1/0/1**" and press "**Enter**" (the blade in slot 1 of the chassis)
    c. Type "**shut**" and press "**Enter**"
    d. Type "**end**" and press "**Enter**"
    e. Type "**write**" and press "**Enter**"

Notice the ping as you shutdown connection to ESM 1. The ping will stop for a few seconds then it will resume until you type the above commands on the second switch. Once you type the above commands on the second switch the ping will stop altogether (no remaining path to the JS12/JS22). It will not resume until you turn one of the switches upstream connection back on with the following commands:
    a. Type "**config t**" and press "**Enter**"
    b. Type "**int G1/0/1**" and press "**Enter**" (the blade in slot 1 of the chassis)
    c. Type "**no shut**" and press "**Enter**"
    d. Type "**end**" and press "**Enter**"
    e. Type "**write**" and press "**Enter**"

Notice the ping will resume. This proves the adapter can fail over since ent0 is hard-wired to ESM1 and ent1 is hard-wired to ESM 2. **This configuration may affect SoL, in particular if the top switch goes down or the port to the top NIC goes down.**

Note: Make sure you re-enable any ports shut-down in the above steps once testing is complete.

C.   Testing Adapter Failover from the BNT Switch Modules

There are a number of ways to test the failover feature, including physically pulling the upstream cables or disabling ports at various locations in the path between the BladeCenter chassis and the upstream switches. The following is an example of how to disable the blade facing ports on the BNT Switch Modules:

Note: BNT switches offer two different CLI's. The original and default CLI (AOS CLI) and a newer, optional CLI (isCLI – industry standard CLI). The commands used will vary depending on the currently configured CLI. AOS is a menu driven interface while isCLI is more Cisco-like in its operation. Selecting the CLI can be done as follows:

In AOS CLI, to cause the switch to automatically default to isCLI on next reload:
> */boot/mode iscli*

In AOS, to cause the switch to prompt at login time for the desired CLI (overrides default selection):
> */boot/prompt ena*

In isCLI, to cause the switch to automatically default to AOS CLI on next reload
> *conf t*
> *boot cli-mode aos*

In isCLI, to cause the switch to prompt at login time for the desired CLI (overrides default selection):
> *conf t*
> *boot cli-mode prompt*

The following demonstrates the same procedure from both AOS CLI and isCLI.

1. Setup looping ping from a network attached device (i.e. your laptop) to the blade with the NIB setup (ping –t ipaddress).
2. Proceed with the following steps (Assumes you are logged in, and in the case of isCLI, you are at the Enable (#) prompt)

Steps	AOS CLI	isCLI
Confirm port names and status	/info/link	show info link
The following commands should be typed on both ESMs, starting with ESM 1	/cfg/port 1/disable apply	config t int port 1 shut

Notice the ping as you shutdown connection to ESM 1. The ping will stop for a few seconds then it will resume until you type the above commands on the second switch. Once you type the above commands on the second switch the ping will stop altogether (no remaining path to the JS12/JS22).

3. It will not resume until you turn one of the connections back on with the following commands:

Steps	AOS CLI	isCLI
The following commands should be typed on both ESMs, starting with ESM 1	/cfg/port 1/enable apply	config t int port 1 no shut

Notice the ping will resume. This proves the adapter can fail over since ent0 is hard-wired to ESM1 and ent1 is hard-wired to ESM 2. This configuration may affect SoL, in particular if the top switch goes down or the port to the top NIC goes down.

Note: Make sure you re-enable any ports disabled in the above steps once testing is complete.

D.   Testing Adapter Failover from the AMM

To Test Adapter Failover from the Management Module do the following:
1.   Setup looping ping from a network attached device (i.e. your laptop) to the server with the NIB setup (ping –t ipaddress).
2.   From the AMM under I/O Module Task select "**Admin/Power/Restart**". Now select the switch in Bay 1 then select "**Restart Module(s) and Run Standard Diagnostics**".   Select "**Ok**" on the pop-window confirming your reboot request.
3.   Now watch the looping ping window. Your ping should only be interrupted for a second but the ping will continue which indicates successful fail over.

Note: The same test can be repeated for the ESM in Bay 2.

VIII. Troubleshooting NIB

If the NIB configuration is not failing over properly, it is important to ensure the NICs are working by doing the following:

To verify connectivity to the Core Switch do the following:
- Remove the NIB configuration from the POWER processor-based blades.
- Create an IP Address on the first NIC (ent0) and then ping the default gateway. Remove this interface from ent0. Now repeat the same steps for second NIC (ent1) and then ping the default gateway.

The following commands can be used to determine the status of NIB:

- The **netstat –d ent#| grep Active** or **netstat –v | grep Active** command will identify if the primary or backup adapter is active from AIX.
- The **netstat –all ent# | grep Active** command will identify if the primary or backup adapter is active from VIOS.

- The **lsdev –Cc adapter** and **lsdev –Cc if** commands will identify ent2 as EtherChannel and en2 as the IP Interface. The **lsattr –El ent2** command will show the current attribute values for the pseudo-device ent2.
- If an adapter failure is triggered by the ping test, an error is posted in the error log. The **netstat –v ent2** command will show information about the primary and backup adapter.

A.   Troubleshooting issues with Ethernet Switch Modules

The following are some of the things you can check to help you resolve issues with ESMs:

- Enable the External Ports for all of the I/O Modules have been enabled from the AMM to ensure connectivity outside of the Chassis.

- To manage the ESM from the Switch make sure you "**Enable**" the "**External management over all ports**". Also make sure you preserve the IP configuration on reset.

- Ensure the downstream ports are not configured for EtherChannel when configuring NIB. When configuring LA ensure the downstream ports are configured for EtherChannel.

- The following table lists some helpful commands for Cisco and BNT switches:

Comment	Cisco IOS CLI	BNT AOS CLI	BNT AOS isCLI
Show running config	*show run*	*/cfg/dump*	*show run*
Dump of information for use by technical support teams	*show tech*	*/maint/tsdmp*	*show tech*
Show information on VLANs	*show vlan*	*/info/l2/vlan*	*show vlan*
Show status of all ports	*show int status*	*/info/link*	*show int link*
Show information about any aggregations	*show eth sum*	*/info/l2/trunk* */info/l2/lacp/dump*	*show portchannel info* *show lacp info*
Show information about individual port status (example port 1)	*show int g1/0/1*	*/stats/port 1/if* */stats/port 1/ether*	*show int port 1* * interface-counters* *show int port 1* * ethernet-counters* (Note: commands above wrapped)

References

IBM BladeCenter JS12 and JS22 Implementation Guide -
http://w3.itso.ibm.com/redpieces/abstracts/sg247655.html?Open

IBM BladeCenter Products and Technology Guide
-http://www.redbooks.ibm.com/redpieces/pdfs/sg247523.pdf

Configuring Network Interface Backup -
- ❖ AIX 5.3 http://publib.boulder.ibm.com/infocenter/pseries/v5r3/index.jsp?topic=/com.ibm.aix.doc/doc/base/aixinformation.htm
- ❖ AIX 6.1 http://publib.boulder.ibm.com/infocenter/systems/scope/aix/index.jsp

Configuring Protected Mode -
http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101034

Cisco Catalyst Switches - http://www-947.ibm.com/systems/support/supportsite.wss/docdisplay?brandind=5000008&lndocid=MIGR-5075938

Ethernet Switch Modules - http://www-947.ibm.com/systems/support/supportsite.wss/selectproduct?taskind=7&brandind=5000020&familyind=5297403&typeind=0&modelind=0&osind=0&psid=bm&continue.x=1

BNT switch documentation - http://www.bladenetwork.net/IBM-BladeCenter-support.html

Forum for asking questions about BNT switches -
http://www.bladeswitching.com

Configuring EtherChannel Bonding in Red Hat -
http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5/html/Deployment_Guide/s2-networkscripts-interfaces-chan.html

Configuring EtherChannel Bonding in SLES -
http://www.novell.com/support/php/search.do?cmd=displayKC&docType=ex&bbid=TSEBB_1222707479531&url=&stateId=0 0 34017274&dialogID=34013800&docTypeID=DT_TID_1_1&externalId=3929220&sliceId=2&rfId=